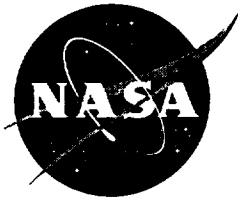# Flight Deck Data Link Displays: An Evaluation of Textual and Graphical Implementations

Alison McGann, Sandy Lozito, and Kevin Corker
Ames Research Center, Moffett Field, California

November 2001

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program Office plays a key part in helping NASA maintain this important role.

The NASA STI Program Office is operated by Langley Research Center, the lead center for NASA's scientific and technical information. The NASA STI Program Office provides access to the NASA STI Database, the largest collection of aeronautical and space science STI in the world. The Program Office is also NASA's institutional mechanism for disseminating the results of its research and development activities. These results are published by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers, but having less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services that complement the STI Program Office's diverse offerings include creating custom thesauri, building customized databases, organizing and publishing research results ... even providing videos.

For more information about the NASA STI Program Office, see the following:

- Access the NASA STI Program Home Page at *http://www.sti.nasa.gov*

- E-mail your question via the Internet to help@sti.nasa.gov

- Fax your question to the NASA STI Help Desk at (301) 621-0134

- Telephone the NASA STI Help Desk at (301) 621-0390

- Write to:
  NASA STI Help Desk
  NASA Center for AeroSpace
  Information
  7121 Standard Drive
  Han..ver, MD 21076-1320

NASA/TM—2001-211384
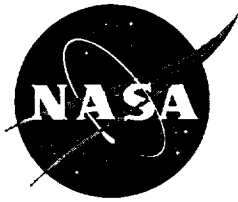
Flight Deck Data Link Displays: An Evaluation
of Textual and Graphical Implementations

Alison McGann
San Jose State University
San Jose, California

Sandy Lozito
Ames Research Center
Moffett Field, California

Kevin Corker
San Jose State University
San Jose, California

November 2001

Acknowledgments

Available from:

Summary

In Experiment 1, 16 pilots participated in a part-task simulation study that evaluated pilot data link communication for short and long message types and for two textual formats. No differences were found between the two textual formats when evaluating data link transaction times and pilot performance on a secondary task. Pilots initiated flight changes more quickly with the T-Scan format, where location of clearance information roughly corresponded to the cockpit instrument layout. Longer messages were less problematic than two short messages sent in close succession as pilots required more verbal clarification for closely spaced messages.

Twenty-four pilots participated in a second experiment that evaluated pilot communication performance for textual data link, two implementations of graphical data link, and a combined graphical and textual information modality. The two modalities incorporating text resulted in significantly faster transaction times and better performance on the secondary task than the two graphical-only implementations. The interval between messages was also more systematically varied in Experiment 2, and a short interval between messages significantly increased the access time for the second message. This delay in access was long enough to increase significantly the total transaction time of the second message, and this effect was exaggerated for the graphical-only implementations. Time to view the message before acknowledgement and time to initiate flight changes were not affected by the interval manipulation. This suggests that pilots adopt a sequential message handling strategy, and presenting messages closely in succession may present operational problems in a data link Air Traffic Control (ATC) environment. The results of this study also indicate that the perceived importance of message content is currently a crucial element in pilot data link communication.

Introduction

Research in air-ground information transfer in the National Airspace System (NAS) has identified several limitations of the current voice communication system. An examination of over 14,000 reports from the Aviation Safety Reporting System (ASRS) by Lee and Lozito (ref. 1) found that as many as 25% cited problems involved air-ground information transfer. The problems cited included inaccurate readbacks, mishears, and misunderstandings between the pilot and controller. In addition, current voice communication channels are often saturated, particularly in busier terminal areas, resulting in frequency congestion.

Digitized information transfer, or data link, should offer several benefits over conventional voice communications. Some of these benefits include a reduction in frequency congestion and a reduction in miscommunications between pilots and controllers. Data link will also allow for permanent storage of information, permitting review of data by the pilots and controllers. In addition, the use of digitized communications will enable technologies permitting more efficient ATC and flight deck automation integration. In fact, data link technology is a central component in the Federal Administration Aviation's (FAA) Air Traffic Control modernization plan.

With the introduction of an air/ground link for digitized information transfer, the communication load between air and ground is likely to shift from an auditory to a visual medium. This shift in primary

modality and in the structure of air/ground communication will introduce new challenges for pilots and controllers. Human computer interface issues, such as display location and data input techniques, are integral to the success of data link as an efficient and safe means of information transfer.

It is likely that the format of data link communications will have an important impact upon its effective use by both pilots and controllers. Due to bandwidth constraints and issues surrounding retrofit cost complexity of extant aircraft, early implementation of data link will likely rely upon the use of alphanumeric textual displays of information. However, as technologies associated with digitized information transfer become more advanced, graphical implementations of data link communications may become more common. Because the cockpit is already a visually loaded environment, the format of data link presentation will be a critical determinant to its effective use as a communication medium. The effective design and implementation of both graphical and textual presentation formats could lead to a reduction in "head-down" time that will be necessitated by a visual form of cockpit data link.

Textual formatting of information will be the mode of introduction for all data link technologies, and its acceptance is contingent in part on an appropriate presentation format. There are specific concerns related to the use of textual information for clearance transmission to the flight deck. Clearance information displayed textually will require greater visual processing time to read and understand the message than traditional voice communication. Previous research has demonstrated that transaction times for textual data link messages are typically twice as long as comparable voice transaction times (ref. 2). Requirements

for the most efficient display of textual communication from the perspective of flight crews need to be investigated. The use of consistent positional references, for example, requiring altitude data to appear consistently on the top line of text in any ATC message, can assist the pilot in visually identifying and referencing those data (ref. 3). In addition, that consistency can provide for a shorter visual scan time to obtain the information.

Graphical display of clearance information could potentially reduce visual processing time as well. In comparing icons to textual information on computer displays, Camacho, Steiner, and Berson (ref. 4) found that icons allowed for more rapid processing. However, LaBerge (ref. 5) reported that individuals require more focal attention to process symbols which are not familiar to them. Wickens (ref. 6) found that pictures were understood at least as well as words if the symbols were familiar to the participants.

The results above suggest that the interface qualities of graphical data link will be critical to its viability as an effective communication medium. One of the concerns associated with a graphical display interface is the location of graphics relative to other display information. The location of data can result in clutter if placed too close to other information on the display. However, if the graphical depictions are too widely dispersed, additional scan and search time by the individual is required. The effects of scanning distance and visual clutter essentially trade off with one another as target dispersion is varied (ref. 6).

Task requirements should also determine the best location for graphical data. If graphical directives are used that offer trend information relevant to an existing display (e.g., an arrow indicating up or down on an altitude display), then close

2

proximity of those displayed data may enhance the pilot's mental model of the system being affected (ref. 6). The *proximity compatibility principle* (PCP) (ref. 7), states that two pieces of information that need to be integrated by the user should be placed in close spatial proximity. Task proximity, or the extent to which two or more sources of information are used as part of the same task, is referenced by the PCP as a critical feature associated with displayed data.

The application of the PCP to flight deck design and data link technology could offer guidelines on the placement of functionally-related information with respect to ATC clearance data. For example, trend information placed next to the associated flight deck display is one application of the PCP to display design. In addition, data link technology for ATC clearance transmission requires the capability for two-way transmission between the ground and the cockpit. Thus, aircraft should be equipped with the ability to receive and transmit air-ground communications, allowing for acknowledgements to instructions from air traffic (e.g., "accept", "reject", and "standby"). The PCP indicates that it may be beneficial to have this information in close proximity to the display of the message and other relevant displays, since the acknowledgement to the message is functionally related to those data. Another display strategy that offers several potential benefits is the use of a redundant information format incorporating both textual and graphical information. A redundant information format allows for the flexibility of users' preferences, as well as allowing for data to be represented in the mode is most salient for that particular information (ref. 6). Also, and more important, redundancy can speeu information processing and reduce errors. In an investigation exploring the comprehension of airline safety cards by passengers, Schmidt and Kysor (ref. 8) compared text, diagrams, and combinations of text with diagrams. They discovered that the cards using mostly words were the least understood, those with only diagrams fared better, and those cards in which words and diagrams were integrated were the best formats for conveying the safety information. Booher (ref. 9) also compared pictorial, verbal, and combined instructional formats, and found the pictorial format with redundant print yielded the best overall performance. Some research suggests that the use of graphical information without text may be detrimental. Brems and Whitten (ref. 10) cautioned against the use of icons without reinforcement by redundant verbal labels due to issues of legibility. However, one potential drawback to redundant information is that more space is occupied, resulting in a more cluttered display (ref. 6).

In an experiment conducted by Hahn and Hansman (ref. 11), the relative effectiveness of voice communication was compared to textual and graphical data link information displays. Error detection of clearance amendment delivery was compared in the three modes in a part-task simulation environment. The data suggested that the graphical implementation resulted in the detection of more erroneous clearances, as well as faster detection of those errors. Additionally, pilots' comments reflected a preference for graphical information due to quickness of evaluation. However, the textual format was preferred for its compactness and accuracy. Hahn and Hansman (ref. 11) suggest that the simultaneous presentation of graphical and textual information should combine the relative advantages of both modalities and they report that the majority of their pilots desired this combined mode of communication.

In evaluating data link as a viable communication medium, issues related to message length and timing should also be considered. Longer messages have been associated with more communication problems in the current voice environment, and this has generated recommendations that controllers reduce the length of their messages (refs. 12, 13). Whether or not data link is similarly affected by message length is an empirical question. Moreover, such reductions in message length assuredly would increase the number of messages that must be delivered in a short period of time. Issues related to message timing also become highly relevant as data link is considered for implementation in the terminal area, an area of dynamic traffic and communication. Issues of congestion and dynamic message handling in either voice or data link will become more pressing in the future because air traffic levels and communication demands are expected to increase. Evaluating data link formats while varying message lengths and timing should reveal potential weaknesses more easily and provide a more representative evaluation of communication performance.

The present studies evaluated textual, graphical, and combined textual and graphical information cockpit display formats as media for conveying data link clearance directives. Because early implementation of data link communication will be limited to textual information transmission, due to bandwidth constraints and retrofit issues for extant aircraft, Experiment 1 evaluated two textual display formats in a part-task simulation. In order to investigate issues relevant to future implementations of data link, Experiment 2 compared textual, graphical and combined information display formats.

Experiment 1

McGann, Lozito, and Corker (ref. 14) examined the impact of different data link textual formats in a simulated cockpit environment. Four alphanumeric formats of air traffic control clearances were compared in a part-task simulator. There were three different accuracy measures: recognition, comparison, and recall of the clearance messages. In addition, reaction time in responding to the messages was measured. No significant differences in accuracy or reaction time were found among the four formats.

The above findings should be interpreted with caution. While it is possible that different textual formats have no effect on pilot performance, it is also possible that differences that may exist were not captured by the measures used in the investigation. All pilots performed quite well on the accuracy and reaction time tasks, with very little difference among the four formats. The participants were commercial pilots and therefore were experts at air/ground communication due to their aviation experience. Hence, it was difficult to induce errors in these tasks, and it is likely that a ceiling effect contributed to the lack of significant performance differences across the formats.

The current study is an attempt to revisit the differences in alphanumeric textual formats in a somewhat different part-task simulation environment. In the previous study, data link clearances were administered out of context; that is, the pilots did not perform a flying task while they received and responded to the clearance data. Since this may affect the task of communicating, this investigation incorporated a manual flying task.

Another difference is the use of a secondary monitoring task as the primary measure of performance differences. It is

4

hypothesized that the format which is easiest to read and interpret will result in better performance on the secondary task as the pilot will have more attention to devote to the secondary activity. In the present investigation, a secondary monitoring task was utilized in conjunction with the flying and communication tasks. The data from the secondary task were intended to provide a more sensitive dependent measure reflecting differences between the textual formats. The use of a secondary measure as a means of estimating reserve capacity and to compare and evaluate visual displays has been well documented (ref. 15).

Format of presentation and message length were the two independent variables manipulated in this study. Two formats, the T-Scan and the Standard format, were compared. A T-Scan format is roughly similar to a cockpit T-Scan layout with consistent positional references, and a Standard format is designed for the next generation aircraft with text elements aligned to the left and a consistent relative order of information. These formats are more fully defined in the Method Section. Because the T-Scan alignment of information is consistent with pilots' previous experience of a T-Scan cockpit layout, and because of the consistent positional references for the T-Scan format, it was hypothesized that the T-Scan format would result in better performance measures. Message length was also varied. Clearances were delivered as either one long message (four commands) or as a pair of short messages (two commands each).

Hypotheses for format differences include the following: 1) Secondary task accuracy will be significantly higher for the T-Scan format; 2) Secondary task reaction time will be significantly lower for the T-Scan format; and 3) Data link acknowledgement times will be significantly faster for the T-Scan format. 4) Finally, it is hypothesized that pilots will initiate flight changes more quickly with the T-Scan format. More specifically, because it should be easier to locate specific clearance directives (i.e. heading or speed) within the T-Scan format, these directives will be input more quickly into the MCP.

Hypotheses for message length include: 1) Secondary task accuracy will be significantly better following shorter messages than longer messages; 2) Secondary task reaction time will be significantly lower following shorter messages than longer messages; 3) Shorter messages will be acknowledged significantly more quickly than longer messages.

Method

Participants

Nineteen current airline pilots were paid to participate in the experiment. The data for three pilots were dropped due to simulator malfunctions. The remaining 16 pilots had a mean age of 41.3 years, ranging from 30 to 57 years. Their average total flight time was 7,175 hours, ranging from 3,000 to 17,000 hours. All of the pilots were type-certified on an advanced transport aircraft (e.g. Boeing 737-300, 757 and 767 aircraft equipped with advanced flight management systems [FMS] and CRT-based display capability).

Equipment

Two Silicon Graphics Personal Iris workstations were used to simulate both an ATC radar display and a glass cockpit flight deck display. The ATC display was equipped with a San Francisco Bay TRACON database and was networked to the pilot workstation, enabling the controller to track the pilot's aircraft on the radar screen and send data link messages

to the flight deck display real time. All data link messages were accompanied by a single aural chime. The two workstations were also connected by a radio-telephone system so that the pilot and controller could talk to each other through conventional means over the radio during the flights.

The pilot display consisted of a primary flight display with speed and altitude strips, an attitude indicator and a moving map display. A portion of the screen was also reserved for data link messages. Pilots controlled thrust by a mouse, and a joystick controlled pitch, yaw, and roll. There was no autopilot capability, so manual control was required throughout the flight. The keyboard was used to enter frequency changes and to acknowledge data link messages. The pilots' primary task consisted of several flights in which they were vectored between San Francisco (SFO) and Sacramento Metropolitan (SMF) airports.

In addition, a Macintosh computer with a touch screen display was positioned next to the flight display and presented the secondary monitoring task for the duration of each flight. For the monitoring task, pilots were asked to monitor four moving gauges and to press an 'event' button on the touch screen each time two or more gauges entered their critical region (Figure 2). Both the gauges and their critical regions fluctuated randomly, making the onset of an event unpredictable and requiring constant monitoring to perform well on this task. Events were programmed to occur approximately three times per minute. The window of opportunity to detect each event was four seconds, and the display would reset itself if the pilot did not respond within this time frame. Monitoring accuracy and reaction time from the event onset were recorded.
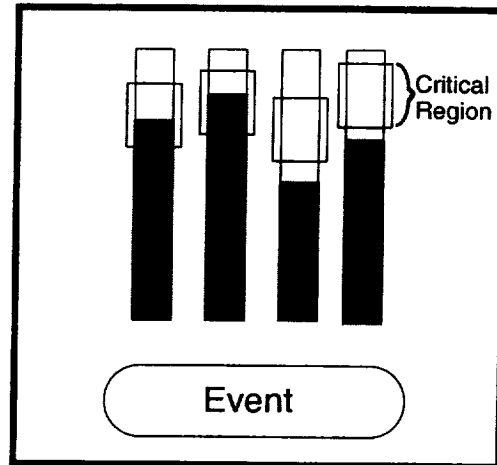

Figure 2. Secondary monitoring task.

Data Link Functionality

Messages sent by ATC automatically appeared on the pilots data link display. There was no step required by the pilot to access a message. Pilots acknowledged to ATC by pressing a key labeled 'accept' on the keyboard.

Instructions and Training

Pilots were first given an overview of the experiment and data link communication. The pilots practiced the secondary task for 2-3 minutes, then flew three practice legs to acquaint themselves with the simulator. The second practice leg was flown with clearances delivered via data link, and the final leg was flown while monitoring the secondary task. During the course of training, pilots received short and long messages in both data link format types used in the experiment.

The pilots were instructed that while performance on the monitoring task was important, they should give the highest priority to flying, then to communicating with ATC, and finally to the secondary monitoring task. They were informed that a 'live' controller was always available on the frequency should they need any clarification. The pilots were instructed that they should accept or reject each data link message within 30 seconds. Messages

not acknowledged within 30 seconds were resent by the controller. Additionally, pilots had the option of clearing data link messages from the display at any time. It was explained that because of development constraints a new message would overwrite the previous message if the display had not been cleared beforehand.

Design

Data link format type and message length were investigated in this study. A related study examined message length in a radio environment (ref. 13). Both studies varied ATC message length with all messages delivered either as one long message (four commands) or as a pair of two short messages in succession (two commands each). For this investigation, each pilot participated with both communication media (voice and data link) and two data link format types (T-Scan and Standard, see Figures 3 and 4). Each of the four legs had its own unique set of scripted clearances. Half of the pilots began with the two voice legs while the other half began with the two data link legs. Within the data link communication modality, the origin airport and the order of format presentation were also varied. This resulted in an experimental design in which voice and data link legs were blocked with the order of communication modality, the order of display formats and order of routes counterbalanced across participants.

Experimental Procedure

The four flight legs flown in this experiment included two from Sacramento Metropolitan to San Francisco International (SMF-SFO) and two return flights (SFO-SMF). Each leg was about 20 minutes in duration and scripted with a unique set of clearances by a retired TRACON controller. Because the simulation had no navigational aids, pilots were vectored the entire route by ATC. Each flight began at about 400 feet on departure and ended after the approach

controller handed the subject aircraft off to the tower frequency. A separate data link format was used for each of the two data link legs. Throughout the flights, long (4 element) messages were alternated with pairs of short (2 element) messages. The short messages were sent roughly 30 seconds apart. There were 18 messages scripted for each leg, consisting of six long and six pairs of short messages. A controller transmitted messages from his workstation keyboard and was also available to respond via radio to any pilot requests.

Two formats were under consideration in this study. One format (T-Scan) used positional encoding and incorporated the layout of the instrument panel (for speed, heading, and altitude). T-Scan is a term commonly used by pilots to represent the "T" shape of the instrument panel configuration. Hence, speed was consistently placed in the upper left, altitude in the upper right, heading in the middle portion of the display, etc. If a particular element was not present in the clearance, its assigned position on the display remained blank (Figure 3). The other format under consideration (Standard) was a proposed format for a next generation aircraft (Figure 4). In this format, all information was aligned on the left and was stacked according to a consistent relative order but did not incorporate positional encoding. A larger font was also used for speed, altitude, and heading numeric values (input that could be loaded into the Mode Control Panel (MCP) or Flight Management System (FMS) in an actual aircraft equipped with these devices). Heading changes did not include directional indications (i.e. right or left turns). In addition, each clearance in this format included a title (ATC CLE₄RANCE), a time stamp, and a status indication along the bottom of the display (ATC CLEARANCE UPLINK).
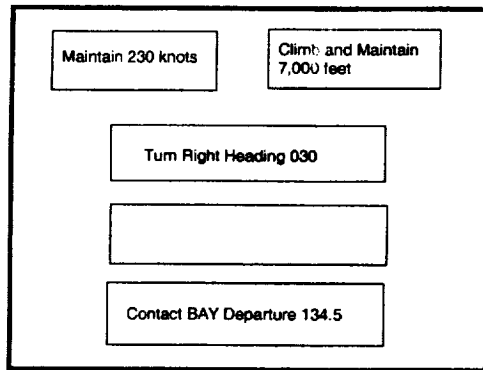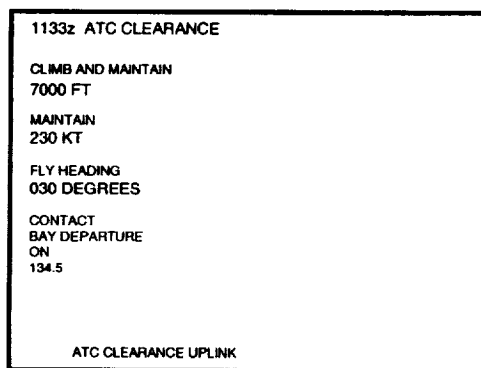
7

Figure 3. T-Scan format.

Figure 4. Standard format.

All information was presented in capital letters for the Standard format versus upper and lower case letters for the T-Scan format.

For the two data link legs, pilots flew vectored routes as instructed by ATC, received all of their messages via data link, and at the same time performed the secondary monitoring task. Two erroneous clearances, one data link message per leg, were included in the communication from ATC. In one such clearance, pilots were asked to reduce speed to 050 knots. In the other, pilots were cleared to land on a non-existent runway (29R at SFO). The rationale for incorporating erroneous clearances ˙ı this study was to give the pilots the option of rejecting a message through voice or data link, and to prevent complacency in accepting data link

messages before they were fully comprehended.

Data from the secondary monitoring task (reaction time and accuracy), the data link transaction times, and responses to erroneous clearances were collected for each of the two data link legs. All voice communication between pilot and controller was recorded on an audio cassette tape. These data allowed us to assess the effects of data link format type (T-Scan or Standard) and message length (two vs. four elements) for six long and six pairs of short messages throughout each flight. Pilots also filled out a questionnaire indicating which format they preferred and which elements of each format they found to be most helpful. The time required for training and data collection for both data link flights was approximately two hours.

Results

Secondary Task Accuracy

Because the pilots were instructed that flying and communicating should be given a higher priority than the secondary task, and because of the modality-specific interference resulting from two visual tasks, secondary monitoring performance during communication was expected to decline. To gain a more precise evaluation of secondary task performance during communication, the secondary events were divided into five time periods after message delivery: 1-10 seconds, 10.5-20 seconds, 20.5-30 seconds, 30.5-40 seconds, and greater than 40 seconds. Each secondary event was also coded by message length (i.e., as having occurred after a long or short message type). There was some variation in the number of events possible per period and per participant that necessitated using the mean accuracy rate per period for each participant in the final analysis. Events occurring immediately

after any of the erroneous clearances were dropped from this analysis.

Secondary task accuracy was then analyzed by a 2 (Format type) x 2 (Message length) x 5 (Period) repeated measures ANOVA. The results indicated no significant main effect for format type. The secondary accuracy rate for the two formats was extremely close with 41% correct for the T-Scan format and 42% correct for the Standard format. Hence, there was no support for the hypothesis that secondary accuracy would be significantly better for the T-Scan format. Significant main effects were found, however, for both message length, F (1,15) = 29.1, p<.001, and period after delivery, F (4,60) = 21.1, p < .001, see figure 5. Thus, secondary task accuracy was significantly better following shorter messages (37% v. 27%). Accuracy also improved significantly as the time between message delivery and event onset increased. A set of pairwise comparisons were performed using a more conservative alpha level (p<.01) and revealed that pilot performance was significantly better for later time periods (greater than 30 seconds and greater than 40 seconds after message delivery) than for earlier time periods.
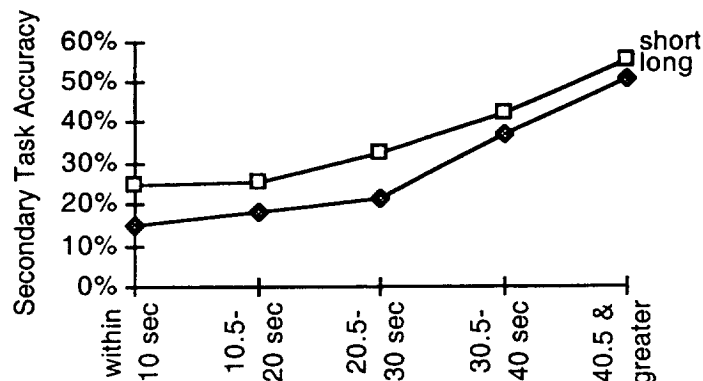
Secondary Task Reaction Time

Because the number of accurate secondary task events was fairly low for the periods closely following message delivery, those periods representing 0-40 seconds were collapsed and compared with period 5 (>40 seconds). This 2 (Format type) x 2 (Message length) x 2 (Period) repeated measures ANOVA revealed no significant main effects or interactions. Thus, there was no support for the hypotheses that format type and message length would affect secondary reaction time. The T-Scan format mean RT was 1.59 seconds (SD=.91) and the Standard format mean RT was 1.58 seconds (SD=.93).

Data Link Acknowledgement Time

Acknowledgement time was defined as the time from message receipt to message acknowledgement. Data link acknowledgement times were analyzed with a 2 (format type) x 3 (message type: first short, second short, and long) x 6 (replication) repeated measures ANOVA.

The replication factor was derived from the number of messages received throughout the flight (6 long and 6 pairs of short messages) and reflects performance over time. There were no significant main effects or interactions for either format type or the replication factor.

Figure 5. Secondary task mean accuracy by length and period after message delivery.
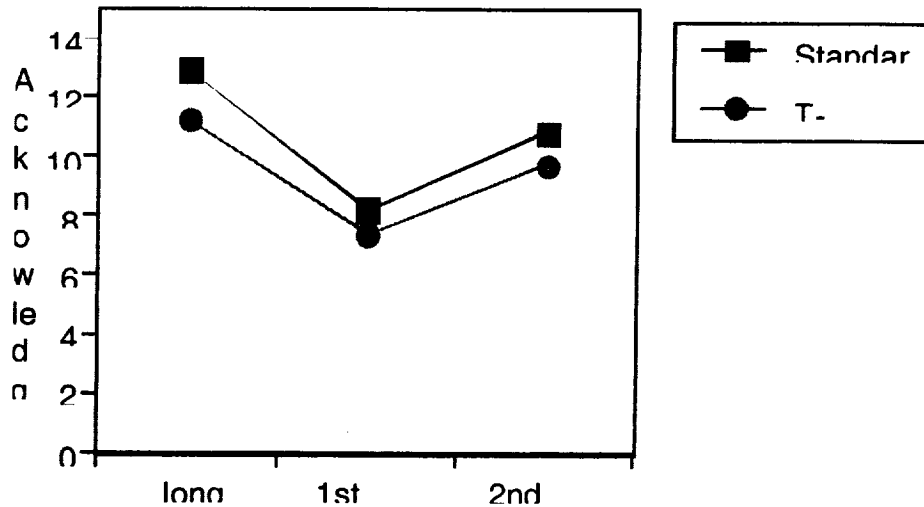
9

Figure 6. Mean data link acknowledgement times for format type and message type.

The overall mean acknowledgement time for the T-Scan format was 9.4 s as compared to 11.0 s for the Standard format (Figure 6). The analysis did reveal a main effect for message type, F (2,30) = 11.99, p<.001. The first short message type had a mean of 7.8 s compared to 10.2 s and 12.1 s for the second short and long message types respectively. Post hoc analyses of simple effects revealed that the first short message type was acknowledged significantly more quickly than either the second short, F (1,15) = 8.75, p<.01 or the long, F (1,15) = 27.75, p<.001 message types.

Erroneous clearances

There were two erroneous clearance one in each of the two data link flight legs. No erroneous clearances were given in the voice radio legs. For data link, one clearance directed the pilots to reduce to 050 knots in cruise, which all of the pilots immediately or eventually rejected. There was no difference based on format type or message length. For the other erroneous message, pilots were cleared for the ILS 29L approach into San Francisco. This approach does not exist, and is in fact a runway used at nearby Oakland Airport.

Two of the sixteen pilots, one for each format type, detected this erroneous runway clearance. Hence, nine out of the 16 erroneous clearances (56%) were immediately or eventually detected by the pilots within each format type. The number of erroneous clearances detected was also distributed equally among short and long messages. The remaining erroneous clearances (44%) were accepted via data link. While the number of erroneous runway clearances not detected by the pilots seems high, it should be remembered that the pilots were flying without their full complement of automation. The part task simulation could not fully simulate the standard crew procedures for an ILS approach (i.e., station tuning and identifying). The high number of erroneous clearances not detected should be interpreted with this in mind.

Verbal Data

Because the pilots occasionally reverted to voice, audio recordings of the verbal communication were transcribed and analyzed separately. These data revealed no significant differences based on format type.
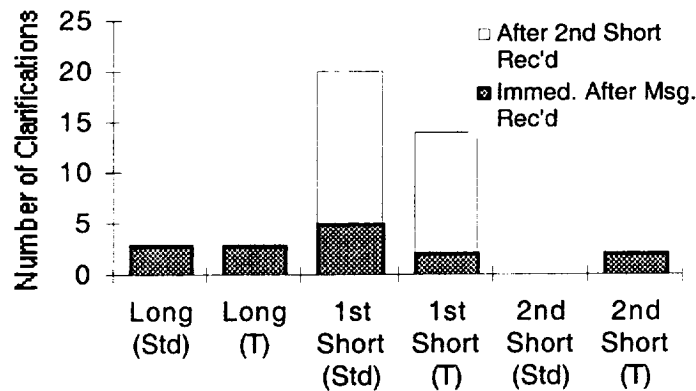
10

Figure 7. Requests for clarification by message type.

Verbal Acknowledgements

Overall, 50 of the 576 messages transmitted (8.7%) were acknowledged verbally for various reasons including need for clarification, checking in on a new frequency verbally, or verbally rejecting the erroneous clearances. For messages acknowledged verbally, 22 or 3.8% of the total message set occurred in the T-Scan format and 28 or 4.9% of the total message set occurred in the standard format. Verbal acknowledgements were not affected by message length. Overall 8.8% of the short messages and 8.3% of the long messages were acknowledged verbally.

Requests for Clarification

Due to the constraints of the data link system used for this study, requests for clarification required the use of the voice channel. Requests for clarification were analyzed in a 2 (format type) x 3 (message type: first short, second short, long) repeated measures ANOVA. This analysis revealed no significant differences based on format type. There were 19 requests for clarification with the T-Scan format, and 23 with the Standard format (clarifications regarding erroneous clearances were not included). Most of these requests concerned direction of turns for heading changes, reasons for being vectored,

whether speed or altitude change was the priority, etc.

There was, however, a main effect for message type, F (2,30) = 13.27, p<.001. As shown in Figure 7, there were more requests for clarification regarding the first short message type, but these requests usually came after the second short message type was already transmitted. (The second short message was transmitted approximately 15 seconds after the first short message was accepted.) It should be noted that the most critical clearances (i.e. heading and altitude changes) were transmitted in the first short message. Additionally, a review log is expected in most data link implementations, and the capability to review previous messages would change the dynamic of this communication.

Time to Initiate Flight Dynamics Changes

Response times for initiating heading and frequency changes were used for these analyses. Because altitude changes were often given before the aircraft had reached the previously assigned altitude, it was impossible to derive response times for these clearances. Similar problems were encountered for airspeed response times. The results for heading and frequency changes must be interpreted with caution
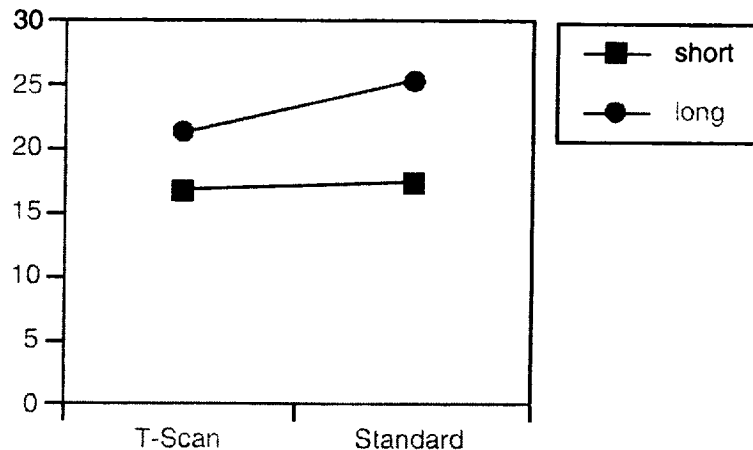
11

Figure 8. Time to initiate heading changes by format and message length.

Table 1. Results Summary.

	T-Scan	Standard	Short	Long
Secondary Accuracy Mean	41%	42%	37%	27%
Secondary RT Mean (s)	1.59	1.58	1.64	1.54
DL Acknowledgement Mean	9.4	11.0	7.8, 11.0	12.6
Erroneous Clearance Detection	9	9	9	9
Number of Verbal Acks.	22	28	34	16
Requests for Clarification	19	23	36	6
Flight Changes	Hdg Faster		Hdg faster, Freq faster	

as the part-task simulation used in this experiment did not replicate a current aircraft type. Nonetheless, the results do indicate some trends.

A 2 (Format type) x 2 (Message length: short, long) repeated measures ANOVA was used to analyze whether heading changes were initiated more quickly based on message length or format type. Message length was reduced to two levels for this analysis because heading changes were always sent in the first short message type. The analysis revealed main effects for format type, F (1,88) = 7.6, p<.01 and message length, F (1,88) = 52.5, p<.001 as well as a two-way interaction, F (1,88) = 3.97, p<.05. Heading changes were

initiated significantly more quickly when presented as part of a short message (17.1 s v. 23.4 s), when presented in the T-Scan format (17.0 s v. 22.7 s), and as seen in Figure 8, the long messages in the Standard format resulted in the slowest input time.

Frequency changes were also analyzed by a 2 (format) x 2 (message length) repeated measures ANOVA. As in the analysis for heading changes, message length was reduced to two levels because frequency changes were always presented in the second short message type. This analysis revealed no main effect or interactions for format type, but a main effect for message length was found, F (1,13) = 11.8, p<.01.
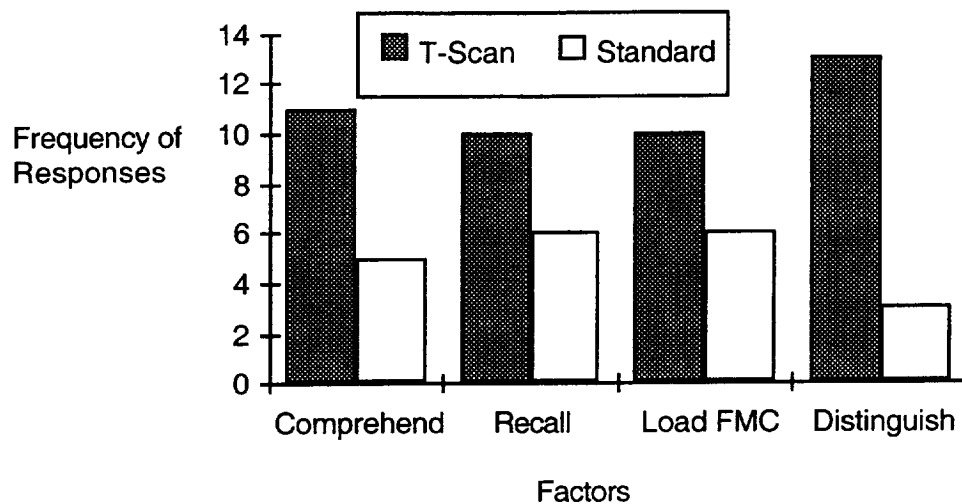
12

Figure 9. Frequency Data for T-scan and Standard format on ease of use for factors.

Subjective Ratings

Frequency changes were input significantly more quickly when presented in a short message then when presented in a long message (16.2 s v. 19.5 s). Table 1 summarizes results for the preceding analyses.

Participants were also given a post experiment questionnaire to record their subjective ratings of the two formats. The pilots were asked which format was easier to read and comprehend, which format was easier to recall, which format would make loading into the flight management computer or mode control panel easier, and for which format it was easier to distinguish the different parts of the clearance. On all four dimensions, the T-Scan format ranked as the most preferred (Figure 9).

The pilots were also asked to rate specific display characteristics as to their usefulness in communicating through data link. High scoring characteristics included the larger font size for critical values as used in the Standard format, positional encoding and upper and lower case alphanumeric characters as used in the T-Scan format.

Additionally, most pilots preferred that display information should be minimized to reduce clutter. Lastly, they tended to prefer abbreviations for feet, knots, and degrees (FT, KTS, and DEG respectively).

Discussion

Formats

The secondary task used was highly sensitive to message length and lag time from message delivery, yet secondary task measures (reaction time and accuracy) did not reveal any significant performance differences based on format type. Thus, if reserve capacity can be taken as a valid measure of information processing, the evidence presented here suggests no advantage to either textual format considered in this study. With the information presented clearly and concisely on a centrally located display, formatting differences had very little effect on pilot performance. It may well be that the benefits of clarity and conciseness gained with the textual implementation of data link mitigate any textual format differences.

13

There was also some evidence that the T-Scan format resulted in faster heading and frequency changes, although caution must be used in interpreting this type of data. Flight parameter changes in a part-task environment are difficult to generalize to actual current aircraft. Neither format resulted in more timely or accurate detection of erroneous clearances, nor did any format require more verbal clarification to supplement the digitized display of information.

It is worth noting that there were some subtle indications of format type differences. Subjective ratings indicated that pilots preferred the T-Scan format across a variety of dimensions. They felt that the T-Scan was easier to comprehend and recall as well as to distinguish among the different clearance directives. They also felt the T-Scan format would allow more simplified direct entry into the mode control panel or flight management system.

Message Length

Significant differences were found based on message length. Secondary monitoring accuracy was significantly better following shorter message types and the type of message also influenced the time for data link acknowledgement. However, it was only the first in each pair of short messages that was acknowledged significantly more quickly (7.8 s) than either the second short (11.0 s) or long message types (12.6 s). Secondary reaction times were not affected by message length. Additionally, there were significantly more verbal requests for clarification of the first short message type, and many of these clarifications occurred after the second short message had been transmitted. Sending messages quickly in succession (within 15 seconds of acknowledging the first short message) may overload pilot cognitive resources and mitigate some of the benefits of data link.

Hence, breaking down longer messages into smaller components may result in longer overall response times. Finally, flight control changes were initiated more quickly with shorter messages for heading and frequency changes.

It should be noted that a similar pattern of results was found in the related voice communication study by Morrow and Rodvold (ref. 13). Pilots were run in a voice only environment with the same message length variables (one long vs. two short message types). For messages delivered via voice, the first short message was also acknowledged significantly more quickly than either the second short or long message types. Additionally, pilots were more likely to request clarification of both the long and first short message types. As in the data link environment, most of the requests for clarification of the first short message type were delayed until after reading back the second short message type. These findings suggest that the communication timing issues uncovered in this investigation are not unique to the data link environment.

In contrast to the data link findings, Morrow and Rodvold also found that the long message types resulted in significantly more voice communication problems than either the first or second short message types. Long messages did not result in more communication problems or verbal acknowledgements in the data link environment and were acknowledged via data link about as quickly as the second short message type. The data link environment may in fact be better equipped to handle long, complex message types (ref. 2) than a voice environment.

Methodology

In relation to the hypothesis that secondary task response time should distinguish the two formats, one

14

methodological consideration in this lack of distinction is suggested. Secondary reaction time did not seem to be a meaningful measure as the window of opportunity to react to each event was short. Secondary task events were programmed to occur rapidly, approximately three times per minute. The time available to react to an event once it occurred was only four seconds, after which time the secondary task reset itself automatically. This short window of opportunity could explain the lack of significant differences based on this measure. Allowing for longer reaction times could have increased the variance for secondary reaction times, but this would have also reduced the difficulty of the secondary task.

Summary

This study found no significant differences based on textual format type. For the data link implementation under investigation, the clear and concise nature of the textual clearance directives was not affected by differences in formatting. However, results from this study suggest that separating or parsing longer data link messages into smaller components has a cost: increased overall response time. The sum transaction time for two short messages put together took longer than one long message by itself. The permanence of the medium in a data link environment (the ability to see and review a message) serves to make it more suitable for longer message types than the voice environment where long messages often overload short term working memory (ref. 13).

Based on the evidence presented here, the authors do not expect that formatting changes to textually-displayed information will significantly improve human performance. Because these textual formats were examined under ideal display conditions, formatting differences could only be expected to result in small differences in performance. Yet an increase in reading speed or comprehension would be worthwhile in a fully integrated environment. It is still possible that poorly formatted information will degrade performance. If the display were more remote or other display characteristics were degraded, the textual formatting would be more critical. While this list is not exhaustive, several guidelines should be kept in mind (ref. 3):

- Each element in the display should be clearly distinguishable;

- Message composition should be concise and compatible with data entry;

- The display should be integrated into a single page or there must be an effective means for identifying and controlling sequential access to different portions of the display.

Variables such as viewing angle, visual acuity, and crew alerting could also impact communication performance and must be taken into consideration in evaluating cockpit data link displays. If these display requirements are adequately addressed, graphical display of information and/or the combined use of synthesized voice may provide incremental performance improvements.

In Experiment 2, several implementations of data link were compared including textual, graphical and combined graphical/textual data link implementations. In addition, because Experiment 1 found that closely spaced messages influenced communication performance, the interval between messages was systematically varied in Experiment 2.

Experiment 2

Experiment 1 found few performance differences in comparing two textual data link formats. With the information presented clearly on a centrally located display, differences in textual display of information had almost no effect on pilot performance. In Experiment 2, the differences between graphical and textual modes as well as a combined graphical and textual mode of conveying data link clearance information were explored. In addition to modality differences, issues relating to time pressure and its effect on data link procedures were investigated.

Because no standard for graphical ATC clearance information has been established, the graphical implementation under investigation in this study was designed to integrate clearance directives and relevant information on the primary flight displays as closely as possible. Thus, a speed directive was placed as closely as possible to the speed indication on the primary flight display. The chosen location of the graphical implementation was intended to help the pilot rapidly assess trend information and the impact of clearance directives on the aircraft flight path. This implementation was designed in accordance with the *proximity compatibility principle* (ref. 7) and borrowed largely from the Hahn and Hansman (ref. 11) data link implementation. An additional graphical data link condition was added in which individual acknowledgements to clearance elements were located in close proximity to the separate elements as displayed on the screen. This enabled us to explore the impact of centralized data link acknowledgements relative to dispersed data link acknowledgements on pilot communication.

Issues associated with message timing have also become more relevant as data link is

being considered for implementation in the terminal area. The dynamic nature of this environment makes message timing factors critical. In light of this, the current investigation also examined communication under time pressure via a manipulation of the interval between data link messages. Experiment 1 found that interval length between messages had an effect on pilot performance. Breaking down long data link messages into pairs of quickly spaced messages resulted in longer overall acknowledgement times for the second message and more need for verbal clarification. It is possible that when clearance messages are presented in rapid succession, interference and disruption of the first message occurs. Moreover, messages occurring in close succession will be likely in an operational environment due to clarifications and amendments of ATC instructions. The authors investigated the effects of time pressure on pilot communication under different modalities.

Additionally, this study examined how the different procedural steps related to ATC message receipt and acknowledgement in the cockpit impacted air-ground communications. Both voice and data link modalities require particular actions in the handling of clearance information. These are somewhat specific to the modality utilized. Data link will require some method of alerting (e.g., visual, aural, or a combination) and will allow for digital acknowledgement as opposed to voice acknowledgement. In contrast to data link, the nature of voice communication is very temporal and compelling. Kerns (ref. 2) posits that a possible advantage of data link as compared to voice communications is that the flightcrew may be able to better manage the communication task around other flight tasks. In this experiment, the procedural steps, and the time it took to complete them, were examined.

As in Experiment 1, a secondary monitoring task was utilized in conjunction with the flying and communication tasks. The secondary task was used to assess differences between the varied conditions. It is hypothesized that the communication modality which is easiest to read and interpret will allow for better performance on the secondary task, as the pilot will have more attention to devote to the secondary activity.

In sum, this study represents an attempt to compare different modes of communicating clearance information (graphics-only, text-only and combined graphics and text), and to determine whether performance in each of these modes is differentially affected by time pressure. Because data link procedures require distinct procedural steps in handling clearance directives, we also attempted to isolate the procedural steps which are most affected by time pressure and mode of presentation. Specifically, we investigated how pilots' message handling was interrupted by the arrival of a new message after a short interval, and whether this depended on the communication modality (graphics-only, text-only or combined). Hypotheses about communication modality and interval follow.

Communication Modality

Because a redundant information format provides users with the flexibility to use either source of information, the graphics and text combined modality should result in better accuracy on the secondary monitoring task as compared to text or graphics alone. Acknowledgement times to ATC should also be reduced for the graphics and text combined modality as compared to graphics and text alone.

Because spatially (and temporally) dispersed acknowledgements force pilots to analyze and acknowledge separate

directives independently, as well as direct them to the source of that information on the aircraft, the graphical condition with spatially dispersed data link acknowledgements should enable more critical evaluation of clearance information than the graphical condition with centralized acknowledgements or the textual condition. This critical evaluation should result in the detection of more erroneous clearances. Additionally, because information is already parsed into distinct units, dispersed acknowledgements will allow for more flexible task scheduling around interruptions, resulting in better accuracy on the secondary monitoring task than the graphical condition with centralized acknowledgements or the textual condition.

Inter-Message Interval

Due to the interruptive nature of the short interval, there should be increased review menu usage for messages separated by short intervals as compared to those separated by longer intervals. Because the time pressure associated with short inter-message intervals will increase pilot workload associated with communication tasks, short communication intervals should result in lower accuracy on the secondary monitoring task as well as longer acknowledgement times than the long communication intervals.

Method

Participants

Twenty-four current airline pilots were paid to participate in the experiment. Twenty-three of the pilots were male, and one was female. The participants had a mean age of 41.5 years, ranging from 30 to 57 years. Their average total flight time was 10,700 hours, ranging from 4,000 to 21,000 hours. All of the pilots were type-certified on an advanced transport aircraft

17

(e.g. Boeing 737-300, 757 and 767 aircraft equipped with flight management systems and CRT-based display capability).

Equipment

The part-task simulation was very similar to that used in Experiment 1, although several upgrades added fidelity to the simulation environment. The displays used to simulate the ATC radar display and the glass cockpit flight display, were upgraded to Silicon Graphics XZ4000 Indigo workstations. As in Experiment 1, the ATC display was equipped with San Francisco Bay and Los Angeles TRACON and Tower databases and was networked to the pilot workstation. This allowed the controller to track the participant aircraft on the radar display as well as send pre-recorded voice and data link messages to the flight deck display in real time. Finally, the two workstations were connected by a radio-telephone system allowing the pilot and controller to talk to each other via the radio during the flights.

For Experiment 2, the flight deck display was now enhanced with a software mode control panel (MCP). The primary flight display and moving map display did not change from Experiment 1. The software-MCP was positioned above the primary flight display, and software knobs on the MCP could be dialed to select values with input from the computer mouse. Participants flew in an autoflight mode selected from the mode control panel. The participants' primary task consisted of several flights in which they were vectored between San Francisco (SFO) and Los Angeles (LAX) airports.

The secondary task was presented via a Macintosh computer with a touch screen display that was positioned next to the flight display. As in Experiment 1, the secondary task was intended to assess the impact of each communication modality

on pilots' "spare capacity". For this visual task, participants were asked to monitor four moving gauges and to press an event button on the touch screen whenever two of the four gauges passed into a critical region (Figure 2). The window of opportunity to detect each event was increased to six seconds, and the display would reset itself if the pilot did not respond within this timeframe. The pre-recorded voice messages were also transmitted through the Macintosh workstation (ref. 16).

Flight data, such as heading and airspeed, MCP entry times and values and data link event times were recorded. All communication between pilot and controller was recorded on audio and video tapes. After the experiment, participants also completed a questionnaire asking them which medium of communication they preferred.

Data Link Functionality

A portion of the pilot's display was also reserved for data link textual messages. Each message was accompanied by an aural alert similar to a selcal chime. A visual alert also appeared at the top of the display (• ATC Message) and remained there until the message was acknowledged (Figures 10 and 11). For the data link conditions, the pilots had to access each message through the new message menu. Unless the pilot was viewing a current message or reviewing a previous message, the data link display defaulted to the new-message menu. New messages could be accessed by clicking on the message bar or new-message menu button with the mouse (Figure 10). New messages were ordered with the most recent on top. One click would open the message for viewing. Once opened, the pilots could accept or reject the message by another mouse click on the accept or reject buttons at the bottom of
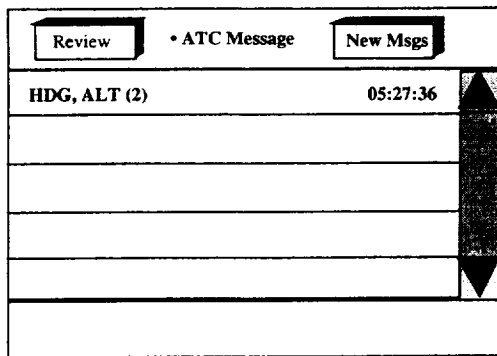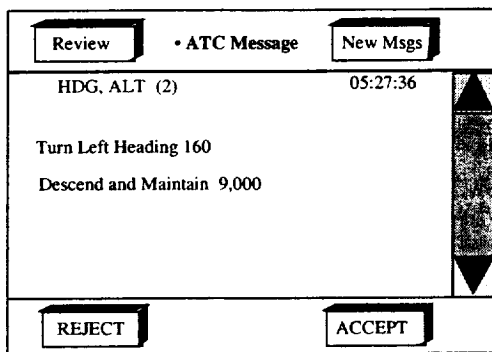
Figure 10. Data link new message menu.

Figure 11. Data link textual clearance.

the textual display (Figure 11). Once accepted, the pilots could enter clearance data directly into the mode control panel by clicking the appropriate knob on the mode control panel display. A single click on each of the speed, heading, and altitude knobs loaded speed, heading, and altitude values respectively. After the current message was accepted, the accept button would change to a clear button and the reject button would disappear. The pilots could then clear the message by either selecting the clear button at the bottom of

the textual display, or selecting the new-message menu button or the review menu button from the menu above the textual data link display. Rejected messages were automatically cleared from the display. All acknowledged messages, either accepted or rejected, could be accessed from the review menu. Messages were ordered in the review menu with the most recent message at the top. Unacknowledged messages, even if they had been opened and viewed, remained in the new message menu until acknowledged. Figure 12 diagrams the discrete procedural steps necessary to comply with a data link clearance.

For the graphical modalities, pilots accessed the data link messages in the same way through the new message menu. Once opened, however, graphical clearance information was displayed directly on the primary flight display. For example, a speed directive was found directly on the speed strip with a transparent green arrow indicating direction and a target value next to the arrow. Once accepted, the graphical icons turned white. When graphical messages were reviewed, information was then displayed in a shadow grey color to prevent confusion and distinguish a reviewed message from the current message. For the first graphical condition, the data link textual display indicated only the type of clearance (i.e. SPD, HDG) and pilots acknowledged the entire message on the lower portion of the textual display just as messages were acknowledged in the textual condition (Figure 13). For the
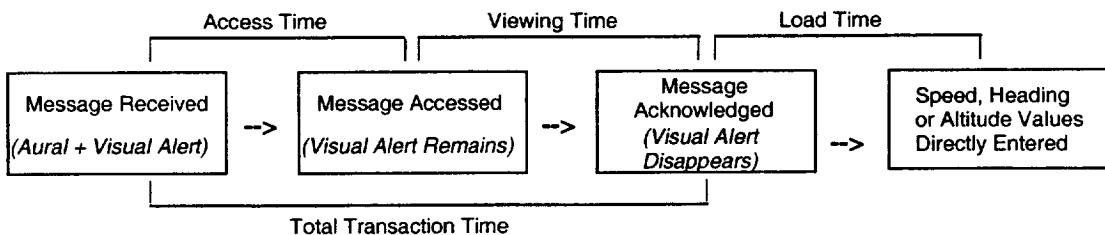

Figure 12. Data link procedural steps.

19

Figure 13. Graphical heading clearance with single acknowledgement.

second graphical condition, no information was contained in the textual display and separate acknowledgements for each directive were co-located directly on the primary flight display. Pilots could then accept one (i.e. speed, heading or altitude) while rejecting the other. Finally, the combined graphical and textual condition functioned just like the textual condition with the addition of redundant graphical information on the primary flight display.

This data link implementation was constrained by the distinct logistical steps required to handle a message. Pilots first had to access the message before they could acknowledge or load information into the MCP. While pilots could manually enter MCP values after the message was accessed, auto-entry of MCP information was precluded in this implementation until the message was accepted. The pilots also had the flexibility to delay any step along the way in order to perform other tasks. While this study represents only one potential data link implementation, most of the constraints associated with it (i.e. message access and acknowledgement) are generic to any implementation.

Instructions and Training

Participants were given an overview of the experiment and data link communication procedures. They were told that the purpose of the study was to investigate the impact of different communication media on air/ground communication. They were not briefed on differences in message interval until after the experiment.

Participants first practiced the secondary monitoring task for 2-3 minutes. After the various aspects of the part-task simulation and its limitations were explained, pilots flew a practice leg (15-20 minutes) while monitoring the secondary task. The same practice route was used for all training legs. Before each data link condition, pilots flew another practice leg with clearances delivered via the medium of their upcoming leg (graphical, textual, or combined). This procedure enabled the pilots to acquaint themselves with data link procedures and ask any questions.

The pilots were instructed that while performance on the monitoring task was important, they should give the highest priority to flying, then to communicating with ATC, and finally to the secondary

20

monitoring task. They were instructed to accept or reject each data link message in as timely manner as they would acknowledge messages in a voice environment.

Procedure

Participants flew a total of six legs, each about 20 minutes in duration. One leg incorporated a purely textual data link implementation, two incorporated different implementations of purely graphical data link, and one incorporated a combined graphical and textual data link implementation. The other two legs (not reported in this paper) examined voice communication and a mixed voice and data link environment. Data from these conditions are presented in another report (ref. 17). All flights began in cruise, and pilots were vectored the entire route by ATC while also monitoring the secondary visual task. Each leg was scripted with a different set of clearances by a retired TRACON controller. The controller transmitted the pre-formatted data link messages from his workstation keyboard and was also available to respond via radio to any pilot requests.

While communication during the cruise phase of each leg consisted of seven or eight messages, analyses focused on two pairs of messages (four critical messages total) embedded in each scenario. Each of the critical messages under investigation contained two elements. The interval between the members of a given pair of

messages in each sequence was varied (either 5 s or 1 min after the pilot accessed the data link clearance, see Figure 14). Non-critical messages were sent between these critical sequences in order to keep the aircraft on course. The order of legs and communication modality were randomized and the order of communication interval was counterbalanced.

The access time, viewing time before acknowledgement, time to enter clearance directives into the MCP, and total transaction time were collected during each of the four legs for each of the four critical messages under investigation. These data allowed us to assess the effects of communication modality (text, graphics, or combined), message interval (5 s or 1 min), and message order (first or second, within each interval) on pilot communication. Secondary task data was also collected throughout all of the flights. The time required for training and data collection across all four flight legs was approximately three hours.

Results

Data Link Procedural Steps

The critical messages under investigation refer to only those pairs of messages where the inter-message interval was systematically varied. Dependent measures included response times for each
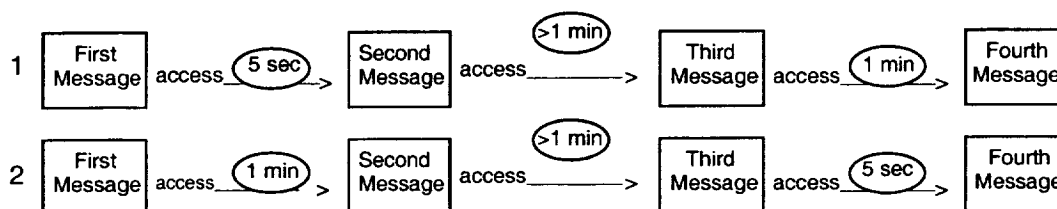

Figure 14. Critical messages under investigation (two counterbalanced conditions).
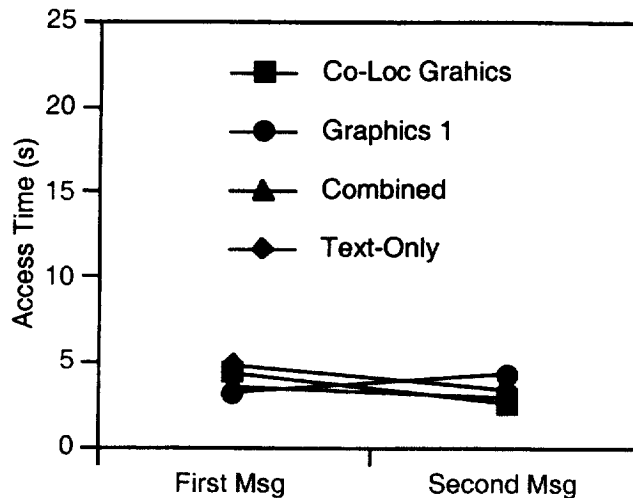
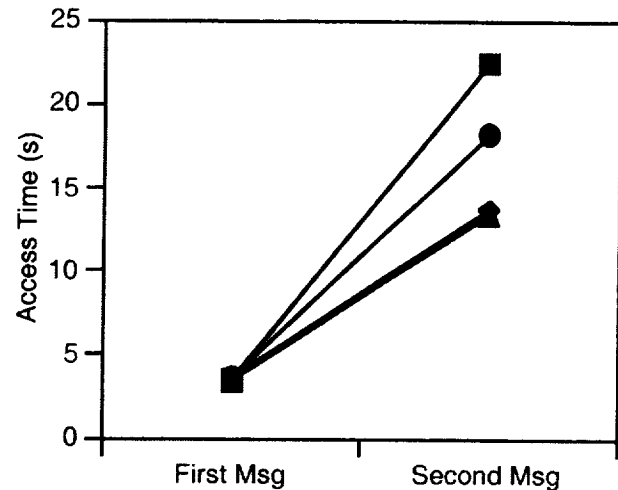Figure 15a. Access time for long-interval messages.

Figure 15b. Access time for short-interval messages.

procedural step required to handle communications via data link. Access time was defined as the time between message receipt and message access (i.e., time to select the new message from the menu). Viewing time accounted for the time between message access and message acknowledgement. Load time was defined as the time from message acceptance to the first direct entry of clearance information into the MCP. Each of these times represents a single and necessary step in the process of handling a data linked ATC clearance (Figure 12). A 4 (Modality) x 2 (Interval) x 2 (Order) repeated measures multiple analysis of variance (MANOVA) was run to analyze the effect of these variables on access time, viewing time and load time. Table 1 summarizes the results of this MANOVA.

Access Time

The analysis revealed a significant 3-way interaction of modality, interval and order on access time, F (3,69) = 6.41, $p<.001$. An interval x order interaction, F (1,23) = 271, p <.001, clearly demonstrates that when data link messages were sent in close succession, access of the second message

in the short interval sequence was significantly delayed relative to other messages. This delay in access suggests that pilots finished handling the first message before they attended to the next incoming message. As seen in Figures 15a and 15b, the 3-way interaction of modality, interval, and order indicates that the delay in access of the second message is exacerbated for the graphics-only modalities when compared to text-only and the combined graphics/text modalities.

Viewing Time

Viewing time accounted for the time between message access and acknowledgement, and the MANOVA revealed an overall main effect of modality on viewing time, F (3,69) = 63.94, p <.001. Because the graphics condition with co-located acknowledgements required two acknowledgement times, it was difficult to compare viewing time for this condition with the other conditions. Viewing time to the first acknowledgement (4.9s) and viewing time to the second acknowledgement (12.3s) are both represented in Figure 16. From the ATC perspective, the second and final
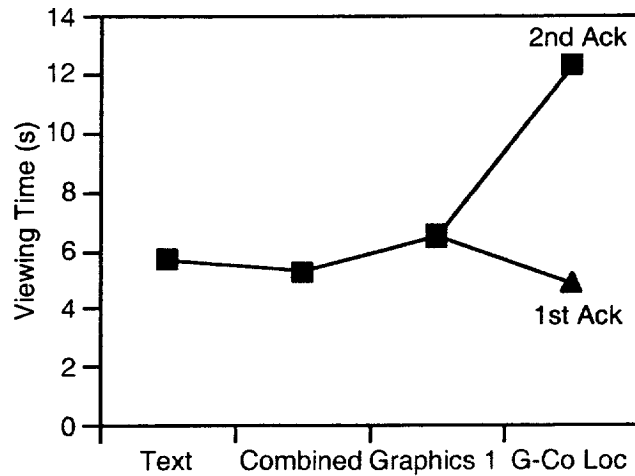
22

Figure 16. Mean viewing time for each modality.

acknowledgement is the operationally relevant measure, so viewing time to the second and final acknowledgement was used for this analysis. As seen in Figure 16, and as revealed in pairwise comparisons, overall viewing time for the graphical condition with co-located acknowledgements was significantly longer than all other modalities, $p<.001$. Pairwise analyses also revealed that the graphics condition with a single acknowledgement was viewed significantly longer than the combined graphics and text condition, $F(1,23) = 9.30$, $p<.01$. For the graphics condition with co-located acknowledgements, the difference in viewing time between the first and second acknowledgement suggests that pilots read and acknowledged each element of the clearance piece by piece rather than reading the whole message before acting.

Load Time

Load time was defined as the time from message acceptance to the first direct entry of clearance information into the MCP. The MANOVA revealed no significant main effects or interactions for load time.

Total Transaction Time

A 4 (modality) x 2 (interval) x 2 (order) repeated measures ANOVA was also run for total transaction time. Total transaction time included time from message receipt to message acknowledgement, with access time and viewing time combined. This measure was included for analysis because it represents the total length of the air-ground dialogue operationally and is comparable with voice communication. The analysis revealed a 3-way interaction of modality, interval, and order, $F(3,69) = 9.43$, $p<.001$, and resulted in the same pattern of findings as was found with access time (Figures 17a and 17b). That is, a short interval between messages significantly increased the transaction time for the second message, and this effect was exaggerated for the graphical modality with co-located acknowledgements. Thus, the effects of modality on viewing time, and the effects of modality, interval and order on access time, were robust enough to lengthen total transaction time.

Additional analyses were run with all the messages included during the cruise portion of the flight (not just the critical message sequences that were part of the interval manipulation). A one-way
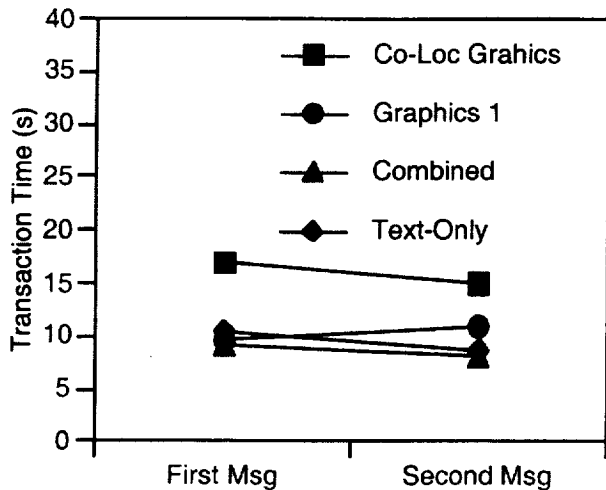
23

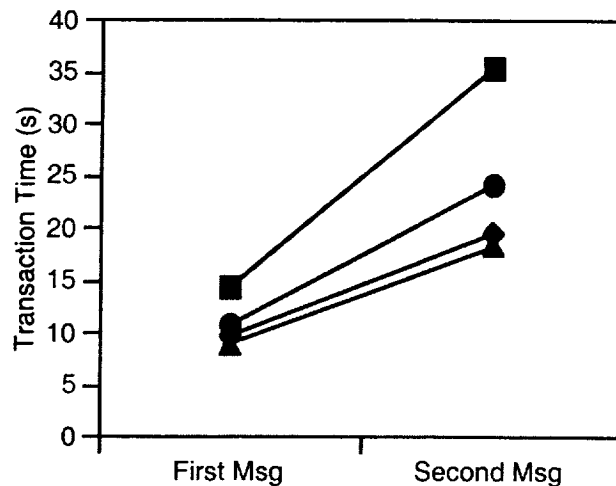Figure 17a. Total transaction time for long-interval messages.

Figure 17b. Total transaction time for short-interval messages.

ANOVA comparing total transaction times for the four modes of communication revealed a significant effect for modality F (3,21) = 27.65, $p<.001$. Pairwise analyses (using an alpha level of $p<.05$) revealed that the combined modality incorporating both graphics and text resulted in siͨﬁcantly faster total transaction times ($10.8s$) than the text-only ($13.3s$) or either of the graphics-only modalities ($15.5s$ and $21.1s$).

Erroneous Clearance Analysis

A one-way ANOVA was run comparing the frequency of detection of erroneous clearances across the four modalities. Since there was only one erroneous clearance per pilot for each modality, power was low for this analyis. The main effect for modality approached significance, F (3,69) = 2.21, $p=.09$. Overall, pilots detected 63% of the erroneous clearances in the text-only condition, 58% in the combined graphics/text condition, 42% in the first graphics-only condition, and 29% in the graphics condition with co-located acknowledgements. It was hypothesizeᴅ that the condition with co-located dispersed acknowledgements would result in more careful evaluation of each

clearance element and greater detection of erroneous clearances. The data do not support this hypothesis. Only two pilots in this condition initially rejected the erroneous element via data link. Five more pilots initially accepted all elements via data link and then later rejected the erroneous portion via the voice channel. Of interest is that for all of the erroneous data link messages rejected across all modalities, 56% were initially accepted via data link and then later rejected through voice communication. Thus, most of the messages were critically evaluated *after* an accept response was sent to ATC.

Secondary Task Analysis

The effect of communication modality on secondary task performance was examined through a one-way analysis of variance. There was some variation in the number of events possible per leg and per pilot which necessitated using the mean accuracy rate per leg for each pilot as the dependent measure in the final analysis. A significant mean accuracy effect for modality was found F (3,69) = 5.39, $p < .01$. The text-only condition had the highest mean accuracy rate at 67% followed by the combined condition (62%), then the graphics condition with co-located

24

acknowledgements (58%) and finally the graphics condition with a single acknowledgement (54%). Pairwise analyses revealed that the text-only modality resulted in significantly better secondary monitoring accuracy than all other modalities except the combined graphics/text modality. No significant main effect was found for secondary task reaction time.

A subset of the secondary task accuracy data was used to assess the impact of inter-message interval on secondary task performance. For this analysis, secondary task accuracy following the second short message was compared with accuracy following the second long message. Thus, only about a quarter of the secondary task events across each leg were included in this analysis. A significant effect for interval was found, F (1,23) = 22.25, $p<.001$, indicating that secondary task performance declined significantly after the second short message (55%) as compared with the second long message (70%). This subset of data did not reveal any modality effect or a modality x interval interaction.

Subjective Ratings

Overall, pilots preferred the combined graphics and text modality (even over voice), although they felt that text alone was easiest to comprehend and recall. Voice communication ranked as the most difficult to recall. The graphics modality with co-located acknowledgements was the least preferred.

Feedback regarding the graphics-only modalities was mixed. Some pilots found the graphical information distracting and felt that it only added clutter to the primary flight display. Other pilots liked the graphical information, but did not like the specific implementation of having arrows placed directly on the altitude and speed strips. Many felt that the graphical

information gave a good overall picture and was easy to interpret. Of those pilots who liked the graphical information, most felt that it was best used as a back-up to textual information.

Pilots were also asked which modality they preferred for short interval sequences: all voice or all data link. Pilots were evenly divided on this issue. Those that preferred a data link environment explained that data link communication allowed them to respond to messages at their 'optimum rate' and on a 'workload permitting' basis. Many also stated that it was a more accurate means of communication. Those that preferred the voice environment cited that voice was more familiar, easier to use, necessitated fewer tasks and allowed for more rapid communication.

Discussion

Modality Differences

The combined graphics/text modality resulted in the fastest data link response times overall while text alone also resulted in significantly faster response times than the graphics-only modalities. Detection of erroneous clearances approached significance in favoring the two modalities incorporating text. Finally, the text-only modality and the combined graphics/text modality resulted in the highest level of performance on the secondary monitoring task. Thus, the primary measures in this study all point to the same trend: text and combined graphics and text resulted in better performance than the graphics-only modalities. Subjective ratings by the pilots indicated that they strongly preferred the combined graphics and text modality.

The two modalities resulting in better performance share the same common denominator of textual information in a precise and concise format. The graphical implementations alone were problematic,

resulting in longer response times due in part to longer scan times (as evidenced by significantly longer viewing times for the graphics-only implementations). Despite the longer scan times, there was no indication that pilots evaluated the graphical clearances any more critically as evidenced in a tendency to detect less of the erroneous clearances. The novelty associated with the graphical depiction of clearance information may partially explain these results. Clearance information on the primary flight display may be initially disruptive. Poor performance and low ratings for the co-located graphical implementation, the most novel, also support the possibility that novelty of implementation influenced the results.

The type of clearances issued may also explain the comparatively lower detection of erroneous clearances and longer response times for the graphical-only implementations. Previous studies demonstrating the advantages of graphical clearances in air/ground communication and more rapid detection of erroneous clearances examined different types of clearance directives. Lee (ref. 18) looked at microburst phenomena while Hahn and Hansman (ref. 11) incorporated erroneous airport destinations. Positional cues are more vital to these clearance types and the advantages demonstrated with a graphical interface were more definitive.

The graphics-only conditions were also more susceptible to delays in access when messages were sent closely in succession. In their subjective evaluations, pilots clearly stated that graphical information should be only be used as a back-up of textual information. Of the two graphics-only conditions, the graphical implementation with the co-located acknowledgements was the least preferred. It resulted in longer overall total transaction times, but at the same time did

not increase vigilance in detecting erroneous clearance directives. In fact, pilots rarely took advantage of the opportunity to reject one portion of the clearance while accepting the other. While the usage in this study suggests no advantage to multiple acknowledgements, this procedure is still very novel and advantages may be clearer as pilots become more accustomed to it. Because the graphical implementations were more problematic than those modalities incorporating textual information, no definitive conclusions regarding multiple acknowledgements can be drawn. It is possible that a multiple acknowledgement interface may be advantageous with textual clearance information, although the impact of this procedure on controllers must also be considered.

However, while the graphical implementations alone did not sufficiently convey clearance intent for speed, heading and altitude changes, graphical clearance information as an enhancement to textual information had no adverse effects on pilot performance. Redundant graphical information allows for both a concise summary and a rapid view of trend information directly next to relevant instrument displays. Despite the fact that graphics without a textual summary increased total transaction times, the combined condition reduced total transaction times as compared to text alone. Additionally, the combined modality did not impede pilots' ability to critically evaluate clearance information as measured by their detection of erroneous clearances. Redundant graphical information was in fact preferred by the majority of the pilots. Having both graphics and textual information allows for individual differences in processing a clearance directive (ref. 6). Graphics and text combined allowed pilots to view clearance directives alongside the display giving a more direct translation of the

26

clearance in relation to the primary flight display while also allowing for a succinct textual summary of clearance directives. The pilots were able to quickly integrate the two sources of information to their advantage.

Procedural Differences

Isolating the individual components of message handling (access time, viewing time, and load time) allowed us to determine how modality differences and time pressure (closely spaced messages) influenced pilots' handling of data link messages. Viewing time was significantly increased for the graphical-only implementations and this contributed to their significantly longer transaction times. Moreover, possibly because of the increased viewing time, the graphical-only implementations were more vulnerable to delays in access when messages were sent in close proximity.

The short inter-message interval only affected the access time of the second message in this experiment. The delay in access was long enough to increase significantly the total transaction time. Thus, as in Experiment 1, Experiment 2 found that messages separated by a short interval have a cost associated with them in increased overall response times. Once a message was accessed, viewing time and load time were not affected by message interval. The delay was isolated to access time, suggesting that pilots can and do schedule when they respond to a data link message. A short inter-message interval did not affect how long they viewed the first message or how long they took to input clearance directives into the mode control panel. Rather it simply delayed their handling of the second message. The second message in the short interval sequence, once opened, was also unaffected by inter-message interval. Time pressure in this study, then, delayed the

initiation of message handling for the second of two successive messages.

This effect of delayed access under conditions of time pressure was found for all modalities under investigation, but it was exaggerated for the graphics-only modalities. Slower modes of communication (in this case the graphics-only modalities) were more heavily affected by short interval sequences than those modalities incorporating text.

Finally, secondary task accuracy was found to decline significantly after the second message of a short interval sequence as compared to the second message of a long interval sequence. While pilots were able to delay their response to data link clearances, decline in secondary task performance after short interval sequences reinforces the notion that this delay was due to limited capacity rather than to any de-emphasis of communication demands. Rapid communication in a data link environment simply left pilots with less capacity to devote to other resources.

Methodological Considerations

The data link and associated graphics under investigation represents only one potential implementation which attempts to integrate clearance trend information onto the primary flight display. Results generated from this implementation cannot be generalized to other graphical implementations or display areas. Moreover, the part-task simulation used represented a single-pilot configuration and did not have a flight management system control display unit. Therefore, the types of clearances under investigation, while realistic, do not encompass the entire range of possibilities. Nonetheless, the results do reveal some germane issues in the development of a graphical data link system from the airborne perspective.

27

Conclusion

The clear and concise nature of textual information can be enhanced with graphics and allow for user preference in incorporating the two sources of information. The results of this study demonstrate the importance of text as a crucial element for data link communication. While the graphical modalities without text resulted in poorer performance measures, graphics and text combined reduced acknowledgement times, and this combination was in fact preferred by the majority of the pilots in this study. Graphical data link may prove to be even more critical with other types of information that rely strongly on positional cues. The two modalities incorporating text (text in a clear and concise format and text enhanced with graphics) allowed for more attention to be devoted to a secondary task, faster acknowledgement times, and more critical evaluation and detection of erroneous clearances. Moreover, these modalities were less susceptible to delays in message handling in rapid communication sequences.

General Summary

The first experiment was designed to assess the effect of format type and message length in textual mode on pilot data link communications. There was no difference between the two formats: T-Scan and Standard format on pilots' performance on secondary task and on datalink transaction times. However, message length had a significant effect on overall response time. It took longer to respond to two short messages than to one long message. But it was found that pilots initiated flight changes more quickly in the T-Scan format. Experiment 2 was meant to change some of methodological issues of Experiment 1 and to evaluate the differences between textual, graphical, and combined textual and graphical modes on datalink communications. The time interval between messages was also manipulated to evaluate the effect of time pressure on pilot datalink communication. It was found that graphical modality when combined with the textual modality had advantages over graphical modality used by itself. The clear and concise nature of text made a difference in the performance of secondary tasks, acknowledgement times, and the detection of erroneous clearances, both in the text only, and combined text and graphical modality. Though the conciseness and clarity of textual mode combined with the graphical data creates redundancy, it has advantages for information processing times and accuracy, which is core to safety in the aviation domain.

References

1. Lee, A.T., & Lozito, S. (1989). Air-ground information transfer in the national airspace system. *Proceedings of the Fifth Symposium on Aviation Psychology.* Columbus, OH.

2. Kerns, K. (1990). *Data Link Communications Between Controllers and Pilots: A Review and Synthesis of the Simulation Literature.* MITRE Corporation Publication MP-90W00027, McLean, VA.

3. Smith, S.L. & Mosier, J.N. (1986). *Guidelines for Designing User Interface Software.* Report No. MTR-10090, ESD-TR-86-278. Bedford, MA: The MITRE Corporation.

4. Camacho, M.J., Steiner, B.A., & Berson, B.L. (1990). Icons versus alphanumerics in pilot-vehicle interfaces. *Proceedings of the 34th annual meeting of the Human Factors Society* (pp. 11-15). Santa Monica, CA: Human Factors Society.

5. LaBerge, D. (1973). Attention and the measurement of perceptual learning. *Memory & Cognition,* 1, 268-276.

6. Wickens, C.D. (1992). *Engineering Psychology and Human Performance* (Second Edition). New York: HarperCollins Publishers.

7. Wickens, C.D., & Andre, A.D. (1990). Proximity compatibility and information display: Effects of color, space and objectness of information integration. *Human Factors,* 32, 61-77.

8. Schmidt, J.K., & Kysor, K.P. (1987). Designing airline passenger safety cards. *Proceedings of the 31st annual meeting of the Human Factors Society* (pp. 51-55). Santa Monica, CA: Human Factors Society.

9. Booher, H.R. (1975). Relative comprehensibility of pictorial information and printed words in proceduralized instructions. *Human Factors,* 17, 266-277.

10. Brems, D.J., & Whitten, W.B. (1987). Learning and preference for icon-based interface. *Proceedings of the 31st annual meeting of the Human Factors Society* (pp. 125-139). Santa Monica, CA: Human Factors Society.

11. Hahn, E.C., & Hansman, J.R. (1992). An experimental study of the effects of automation on pilot situation awareness in the data link ATC environment. *The International Journal of Aviation Psychology.*

12. Cardosi, K.M. (1993). An Analysis of en route controller-pilot communications (Rep. No. DOT/FAA/RD - 93/11). Washington, DC: Department of Transportation, Federal Aviation Administration.

13. Morrow, D.G., & Rodvold, M.A. (1993). *The influence of ATC message length and timing on pilot communication.* NASA Ames Contractor Report 177621.

14. McGann, A., Lozito, S., & Corker, K. (1992, October). Cockpit data link displays: An evaluation of textual formats. *Proceedings of the SAE International Aerotech Conference.* Anaheim, CA.

15. Ogden, G.D., Levine, J.M., & Eisner, E.J. (1979). Measurement of workload by secondary tasks. *Human Factors,* 21(5), 529-548.

29

16. Pisanich, G., Lee, E., & Beck, L. (1994). A part-task simulator for advanced automation and communications research. *Proceedings of the AIAA Flight Simulation Technologies Conference*. Scottsdale, AZ.

17. McGann, A., Morrow, D.G., Rodvold, M.A., & Macintosh, M.A. (1994). *A comparison of pilot communication using a single medium vs. mixed voice and data link ATC environment.* Unpublished manuscript.

18. Lee, A.T. (1991). Aircrew decision-making behavior in hazardous weather avoidance. *Aviation, Space, & Environmental Medicine*, 159-161.

Report Documentation Page	Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE November 2001	3. REPORT TYPE AND DATES COVERED Technical Memorandum

4. TITLE AND SUBTITLE
Flight Deck Data Link Displays: An Evaluation of Textual and Graphical Implementations

6. AUTHOR(S)
Alison McGann, Sandy Lozito, and Kevin Corker

5. FUNDING NUMBERS
NCC 2-798

DTFA01-93-Z-02087

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
NASA Ames Research Center
Moffett Field, California 94035-1000

8. PERFORMING ORGANIATION REPORT NUMBER
IH-023

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
NASA/TM—2001–211384

11. SUPPLEMENTARY NOTES
Point of Contact: Sandra Lozito, M/S 262-4, Ames Research Center, Moffett Field California 94035
(650) 604-0008

12A. DISTRIBUTION/AVAILABILITY STATEMENT
Subject Category: 03-01   Distribution: Public
Availability: NASA CASI (301) 621-0390

12B. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

In Experiment 1, 16 pilots participated in a part-task simulation study that evaluated pilot data link communication for short and long message types and for two textual formats. No differences were found between the two textual formats when evaluating data link transaction times and pilot performance on a secondary task. Pilots initiated flight changes more quickly with the T-Scan format, where location of clearance information roughly corresponded to the cockpit instrument layout. Longer messages were less problematic than two short messages sent in close succession as pilots required more verbal clarification for closely spaced messages.

Twenty-four pilots participated in a second experiment that evaluated pilot communication performance for textual data link, two implementations of graphical data link, and a combined graphical and textual information modality. The two modalities incorporating text resulted in significantly faster transaction times and better performance on the secondary task than the two graphical-only implementations. The interval between messages was also more systematically varied in Experiment 2, and a short interval between messages significantly increased the access time for the second message. This delay in access was long enough to increase significantly the total transaction time of the second message, and this effect was exaggerated for the graphical-only implementations. Time to view the message before acknowledgement and time to initiate flight changes were not affected by the interval manipulation. This suggests that pilots adopt a sequential message handling strategy, and presenting messages closely in succession may present operational problems in a data link Air Traffic Control (ATC) environment. The results of this study also indicate that the perceived importance of message content is currently a crucial element in pilot data link communication.

14. SUBJECT TERMS
Message format, Controller Pilot Data link Communication (CPDLC), Aviation Communication

15. NUMBER OF PAGES
32

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited